Measuring Human Performance in Battle Command

Dr. Carl W. Lickteig, William R. Sanders,
Dr. Paula J. Durlach and Dr. James W. Lussier

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is the Army's lead laboratory conducting research, development and analysis on training, leader development and Soldier issues. ARI's focus is on the entire Soldier life cycle of recruitment, selection, assignment, training, retention and mission performance. ARI provides the behavioral science and technology (S&T) tools to help the Army of the future realize its goals for superior performance across the full spectrum of conflict in all operational environments.

Battle command and human performance during sustained operations in harsh, unpredictable environments is a key focus for ARI. Here, Soldiers from Charlie Co., 2/162d Infantry Regiment perform a quick response force mission in Sadr City, Iraq. (U.S. Air Force photo by SSGT Ashley Brokop.)

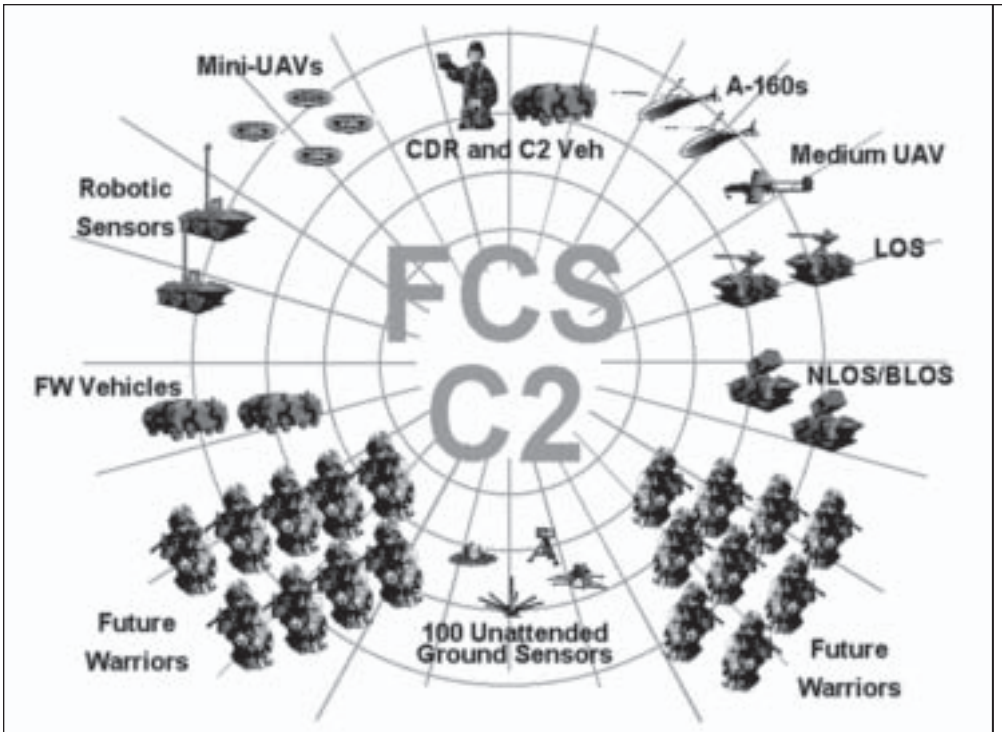
The rapidly evolving strategic environment and continuous advances in technology — particularly the digitization of information systems — demand new approaches to applied research and development (R&D) in the personnel arena as well as in weapon systems development. Factors such as organizational redesign, as represented by the Army Campaign Plan, will also have ripple effects in terms of training strategies at all levels. There are probably no S&T programs that do not have to take these factors into account by simultaneously working to solve today's problems while anticipating tomorrow's challenges. Clearly, the optimum outcome is achieved by implementing a mutually reinforcing balance to answer present and future needs. ARI adopted this approach in many of its programs, of which one prominent example was the measurement of human performance for the Defense Advanced Research Projects Agency (DARPA)/ Communications-Electronics Research, Development and Engineering Center Future Combat Systems Command and Control (FCS C2) Program.

> ARI provides the behavioral science and technology tools to help the Army of the future realize its goals for superior performance across the full spectrum of conflict in all operational environments.

To assess the human performance essential for battle command in FCS, ARI joined the program for an iterative series of commander-in-the-loop experiments. ARI focused on four command group (cmd grp) members — commander, information manager, battlespace manager and effects manager — located in a mock-up C2 vehicle as they planned and executed more than 40 virtually simulated battle runs across 4 experiments. Objective and subjective measurement methods developed by ARI were used to quantify and understand how a future cmd grp might perform the basic battle command functions of plan, move, see and strike.

Experimental Design

To explore new approaches for battle command, the design iteratively raised cmd grp responsibilities and FCS capabilities across experiments based on lessons learned in prior experiments. The design systematically varied battle run complexity (medium, high and too high) within and across experiments by increasing enemy force activity and size, eliminating friendly



FW – Future Warriors
UAV – unmanned aerial vehicle
LOS – line of sight
NLOS – non-line-of-sight
BLOS – beyond-line-of-sight

Figure 1. Organization of the unit cell.

assets and inserting civilians on the battlefield. Complexity was used to gauge the performance limits of the cmd grp members and a notional FCS organization. The design also stressed a deliberate practice approach — the repetition of similar runs with feedback — to ensure experimental results were based on a proficient cmd grp. Typically, the four primary experiments lasted 10 days each, beginning with 3 days of training followed by 10 battle runs with planning and execution phases requiring approximately 60-90 minutes per phase.

Human Performance Results

The FCS concept calls for an unprecedented human-machine alliance that will require warfighters to employ a network-linked force of unmanned systems in concert with manned systems to perform future missions. To measure human performance in future FCS cmd grps, ARI developed objective measures for the cmd grp's verbal communications and human-computer interactions (HCI). Subjective measures were developed to assess key issues including workload, training, system performance and human performance. Efforts were also made to develop and validate automated HCI measures. The construct is depicted in Figure 1.

Verbal communications by the cmd grp in the C2 vehicle were a near-continuous activity, occurring 93 percent of the time during the execution phase of each run. A pattern of steady conversation occurred despite participants' common access to a visually rich and timely battlefield situation depiction on their C2 displays. By far, the majority of communications were devoted to "seeing" the battlefield by collecting and interpreting data from multiple ground and air sensors to construct an accurate battlefield situation understanding.

HCI analysis revealed that more than 1,000 separate HCI actions were typically performed by the cmd grp during the battle run's execution phase. Detailed analysis of HCI data quantified and related the demands placed on each cmd grp member to accomplish the basic plan, move, see and strike battle command functions. The HCI assessment revealed the importance of examining how task demands changed within and across battle runs as well as

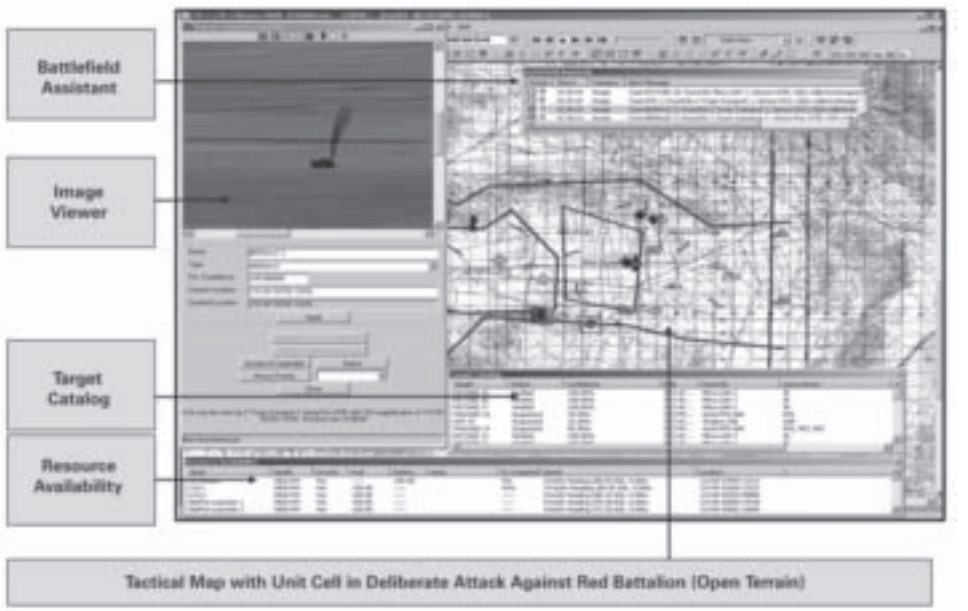


Figure 2. Sample screen view of a C2 prototype interface.

how the cmd grp learned to reallocate tasks across members to cross-level the workload. Figure 2 provides a sample view of the screen information cmd grp members used during the battle runs.

Subjective measures of workload helped assess the FCS goal to reduce the cmd grp's size for a small unit with numerous robotic elements. Moderate to high levels of workload were reported by the information and battlespace managers who directed the actions of multiple sensors while interpreting and sharing the battlefield intelligence obtained. As workload increased at the "too high" level of complexity, the information and battlespace managers' performance ratings sharply declined. System performance ratings by participants were also captured to refine current features and to add new features and more automation to the simulated FCS across experiments.

As workload increased at the "too-high" level of complexity, the information and battlespace managers' performance ratings sharply declined.

Automated measures of HCI are important tools for understanding cmd grp performance requirements. However, the vast amount of HCI data available cannot be readily captured through manual recording methods. Efforts to develop automated HCI measures were only partially successful, underscoring the need for more development to ensure such data are captured for future FCS training and evaluation efforts.

Novice versus expert cmd grp performance was also examined. A key FCS C2 program feature was the commitment to create an expert cmd grp of lieutenant colonels (LTCs) for the four primary experiments to explore future battle command concepts. An excursion experiment was also conducted with a novice cmd grp comprising U.S. Military Academy and Reserve Officers' Training Corps cadets. Comparisons of novice versus expert performance based on the ARI measurement approach found that novices:

- Spent more time in silence, less time collaborating.
- Talked more about firing, less about seeing.
- Talked more about own troops, less about enemy.
- Talked more about enemy location, less about enemy size, type and disposition.
- Performed fewer computer interactions to recognize and identify targets.
- Performed more computer interactions to assess battle damage.

Overall, the novice group seemed to approach their C2 roles with a hasty "find-and-kill" mindset — not unlike a video game mentality — while the more expert cmd grp deliberately strove to build an accurate and complete battlefield situational understanding before engaging the enemy.

Human Performance Issues

The research shed needed light on key human performance issues associated with the introduction of advanced and complex technologies. Some conclusions on workload, training and proactive research are discussed below.

Command group members conduct an after action review in their mock C2 vehicle. (ARI photo.)

Workload is a serious concern with the emergence of numerous unmanned and complex systems, including ground/air sensors and beyond-line-of-sight weapons. Technology may overwhelm the cmd grp's ability to provide robotic and human force C2 while being inundated with potential deluges of battlefield information. Objective and subjective data confirmed increasingly high levels of workload as battle runs became more complex. Increased levels of automation can reduce workload, but they can also increase it.

The ultimate value of a C2 system is determined not so much by technology per se, but by shaping technology to complement human performance.

A pattern of decreasing workload based on objective and subjective data during the first three experiments was attributed to technology — the iterative insertion of new and more automated features across experiments. However, this pattern was reversed in the fourth experiment despite the LTCs' increased expertise and their advanced FCS capabilities. The reversal may reflect a recurrent finding that expectations about doing more with technology often result in greater burdens on warfighters.

Training is the glue that will hold FCS and the Future Force together. FCS

cmd grps will require extraordinary levels of tactical and technical expertise. A small cmd grp with robotic elements must reformulate battle commands into computer commands. Today's succinct verbalizations, such as commander's intent and guidance, entail many implied tasks for humans. In the future, unmanned systems may require directives issued in computer-mediated formats with more explicitly and precisely defined tasks. Furthermore, expertise may be more perishable as C2 system software changes quickly to meet evolving requirements. Even though the LTCs helped design their C2 system, they had difficulty in understanding complex input requirements and the operational consequences of highly automated functions. The LTCs stressed that training was required to provide more hands-on experience in tactical scenarios, more emphasis on employment techniques and more opportunities to devise and ingrain standard operating procedures.

Proactive Research. Historically, Army materiel researchers have had difficulty conducting adequate early assessments of the human dimension in system performance. Human performance is critical for FCS because empowering commanders through advanced C2 systems is at the heart of the FCS concept. The revolutionary nature of the Army's transformation — as embodied in the FCS acquisition program — increases the risk of relying exclusively on traditional assessment methods such as C2 hardware and software component tests or the outcomes of simulation without warfighters-in-the-loop. ARI's measurement methods and the results on human performance provided reliable and empirical data for important and timely decisions on training, materiel, manpower and personnel. Findings were readily transitioned to acquisition efforts through DARPA's dual roles in FCS simulation and acquisition. The FCS C2 program was cited by the FCS Integrated Product Team for Training as a key contributor to their design planning. The human performance findings shaped the C2 prototype showcased in the Capstone Demonstration prior to FCS Milestone B. The Army recognized ARI's contribution to the human dimension of battle command with the 2003 Research and Development Achievement Award.

With respect to the human performance essential to battle command, the ultimate value of C2 R&D programs is determined as much by the investment in training and evaluation as the investment in simulation. The ultimate value of a C2 system is determined not so much by technology per se, but by shaping technology to complement human performance.

Bravo Co., 103rd Armored Battalion's Soldiers patrol the streets of As Siniyah, Iraq, at dawn Jan. 12, 2005. These troops are part of the 28th Infantry Division's Task Force Dragoon. ARI is actively shaping C2 technology to complement battlefield commanders' decision-making capability. (U.S. Army photo by SPC Elizabeth Erste.)

DR. CARL W. LICKTEIG is a Senior Research Psychologist for ARI at Fort Knox, KY, with expertise in training, evaluation, human performance and the application of digital technology. He has a B.A. in social psychology from Bellarmine University and both an M.A. and Ph.D. in experimental psychology from the University of Louisville.

WILLIAM R. SANDERS is a Research Psychologist for ARI at Fort Knox, with expertise in skill acquisition, skill retention and human performance measurement. He received his B.A. in industrial/organizational psychology from the University of California-Berkeley and his M.A. in industrial/organizational psychology from Ohio State University.

DR. PAULA J. DURLACH is a Research Psychologist for ARI in Orlando, FL, with expertise in associative learning, individual differences and training. She earned B.A. degrees in psychology and mathematics from Swarthmore College and both an M.A. and a Ph.D. in psychology from Yale University.

DR. JAMES W. LUSSIER is the Team Leader for ARI at Fort Knox, with expertise in battle command training methods and tactical training of future leaders. He received his B.A. in psychology from Columbia University and both an M.A. and Ph.D. in experimental psychology from Fordham University.